

Docket No.: 49657-850



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of :
Christian KROOS, et al. :
Serial No.: : Group Art Unit:
Filed: November 15, 2000 : Examiner:
For: METHOD AND APPARATUS FOR MEASURING MOTION OF AN OBJECT
SURFACE BY MULTI-RESOLUTION ANALYSIS USING A MESH MODEL

**CLAIM OF PRIORITY AND
TRANSMITTAL OF CERTIFIED PRIORITY DOCUMENT**

Assistant Commissioner for Patents
Washington, DC 20231

Sir:

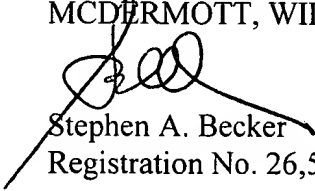
In accordance with the provisions of 35 U.S.C. 119, Applicants hereby claim the priority of:

Japanese Patent Application No. 11-325362,
filed November 16, 1999

cited in the Declaration of the present application. A certified copy is submitted herewith.

Respectfully submitted,

MCDERMOTT, WILL & EMERY


Stephen A. Becker
Registration No. 26,527

600 13th Street, N.W.
Washington, DC 20005-3096
(202) 756-8000 SAB:klm
Date: November 15, 2000
Facsimile: (202) 756-8087

49657-850
NOVEMBER 15, 2000
KROOS et al.

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

McDermott, Will & Emery

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.



出 願 年 月 日

Date of Application:

1999年11月16日

出 願 番 号

Application Number:

平成11年特許願第325362号

出 願 人

Applicant(s):

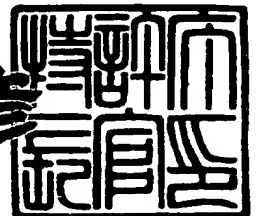
株式会社エイ・ティ・アール人間情報通信研究所



2000年 1月 7日

特許庁長官
Commissioner,
Patent Office

近 藤 隆 彦



出証番号 出証特平11-3092511

【書類名】 特許願

【整理番号】 1991478

【提出日】 平成11年11月16日

【あて先】 特許庁長官殿

【国際特許分類】 G06T 17/00
G06T 9/00

【発明者】

 【住所又は居所】 京都府相楽郡精華町大字乾谷小字三平谷 5 番地 株式会
社エイ・ティ・アール人間情報通信研究所内

 【氏名】 クリスチャン・クロス

【発明者】

 【住所又は居所】 京都府相楽郡精華町大字乾谷小字三平谷 5 番地 株式会
社エイ・ティ・アール人間情報通信研究所内

 【氏名】 エリック・バティキオティス・ベイツン

【発明者】

 【住所又は居所】 京都府相楽郡精華町大字乾谷小字三平谷 5 番地 株式会
社エイ・ティ・アール人間情報通信研究所内

 【氏名】 倉立 尚明

【特許出願人】

 【識別番号】 592179296

 【住所又は居所】 京都府相楽郡精華町大字乾谷小字三平谷 5 番地

 【氏名又は名称】 株式会社エイ・ティ・アール人間情報通信研究所

【代理人】

 【識別番号】 100064746

 【弁理士】

 【氏名又は名称】 深見 久郎

【選任した代理人】

 【識別番号】 100085132

 【弁理士】

【氏名又は名称】 森田 俊雄

【選任した代理人】

【識別番号】 100091395

【弁理士】

【氏名又は名称】 吉田 博由

【選任した代理人】

【識別番号】 100096781

【弁理士】

【氏名又は名称】 堀井 豊

【選任した代理人】

【識別番号】 100096792

【弁理士】

【氏名又は名称】 森下 八郎

【手数料の表示】

【予納台帳番号】 008693

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 対象物の表面動き測定方法および装置、ならびに当該方法を実現するようコンピュータを動作させるためのプログラムを記憶したコンピュータ読取可能な記録媒体

【特許請求の範囲】

【請求項 1】 先行する第 1 の画像フレームおよび前記第 1 の画像フレームより後の第 2 の画像フレームを含む、対象物の画像フレームのシーケンスから、前記対象物の表面の各部位の動きを測定するための表面動き測定方法であって、

相互の隣接関係が定められた複数個のノードを含むメッシュモデルによって前記第 1 の画像フレームの画像内の前記対象物の表面の各部位の位置を表現するステップと、

予め定められた複数通りの解像度で前記第 1 の画像フレームの画像および前記第 2 の画像フレームの画像を各々複数個のサブバンド画像に変換するステップと

、前記メッシュモデルから、前記メッシュモデルのノード密度より低く、かつ互いに異なるノード密度を有する複数個の粗メッシュモデルを生成するステップと、前記複数個の粗メッシュモデルは、それぞれ前記複数個のサブバンド画像と対応付けられており、

前記第 1 の画像フレームおよび前記第 2 の画像フレームの各々の前記複数個のサブバンド画像に対してそれぞれ前記複数個の粗メッシュモデルのうちの対応するものを重ねたオーバーレイ画像を得るステップと、

前記第 1 の画像フレームおよび前記第 2 の画像フレームの、それぞれ対応するサブバンド画像から得られた前記オーバーレイ画像のテクスチャマップのクロス相関を計算することにより、前記第 2 の画像フレームの前記メッシュモデルの各ノードと、前記第 1 の画像フレームの前記メッシュモデルの各ノードとを対応付け、各ノードの前記第 1 の画像フレームから前記第 2 の画像フレームの間の相対的位置変化を算出するステップとを含む、表面動き測定方法。

【請求項 2】 前記算出するステップは、

前記第 1 の画像フレームおよび前記第 2 の画像フレームの、第 1 のサブバンド

画像から得られた前記オーバーレイ画像のテクスチャマップのクロス相関計算を行なうことにより、前記第2の画像フレームの前記第1のサブバンド画像に対応する第1の前記粗メッシュモデルの各ノードと、前記第1の画像フレームの前記第1のサブバンド画像に対応する前記第1の粗メッシュモデルの各ノードとを対応付け、前記第1の粗メッシュモデルの各ノードの前記第1の画像フレームから前記第2の画像フレームの間の相対的位置変化を算出するステップと、

前記第2の画像フレームの、前記第1の粗メッシュモデルよりも高いノード密度を有する第2の粗メッシュモデルに対して、前記第1の粗メッシュモデルの各ノードの位置変化に応じた変形を行なうステップと、前記第2の粗メッシュモデルは、前記第1のサブバンド画像よりも高い空間周波数を有する第2のサブバンド画像に対応するものであり、

前記第1の画像フレームおよび前記第2の画像フレームの、前記第2のサブバンド画像から得られた前記オーバーレイ画像のテクスチャマップのクロス相関計算を行なうことにより、前記第2の画像フレームの前記第1のサブバンド画像に対応する前記第2の粗メッシュモデルの各ノードと、前記第1の画像フレームの前記第2のサブバンド画像に対応する前記第2の粗メッシュモデルの各ノードとを対応付け、前記第2の粗メッシュモデルの各ノードの前記第1の画像フレームから前記第2の画像フレームの間の相対的位置変化を算出するステップとを含む、請求項1に記載の表面動き測定方法。

【請求項3】 先行する第1の画像フレームおよび前記第1の画像フレームより後の第2の画像フレームを含む、対象物の画像フレームのシーケンスから、前記対象物の表面の各部位の動きを測定するための表面動き測定方法を実現するようコンピュータを動作させるためのプログラムを記憶した、コンピュータ読取可能な記録媒体であって、

前記表面動き測定方法は、

相互の隣接関係が定められた複数のノードを含むメッシュモデルによって前記第1の画像フレームの画像内の前記対象物の表面の各部位の位置を表現するステップと、

予め定められた複数通りの解像度で前記第1の画像フレームの画像および前記

第2の画像フレームの画像を各々複数個のサブバンド画像に変換するステップと

前記メッシュモデルから、前記メッシュモデルのノード密度より低く、かつ互いに異なるノード密度を有する複数個の粗メッシュモデルを生成するステップと、前記複数個の粗メッシュモデルは、それぞれ前記複数個のサブバンド画像と対応付けられており、

前記第1の画像フレームおよび前記第2の画像フレームの各々の前記複数個のサブバンド画像に対してそれぞれ前記複数個の粗メッシュモデルのうちの対応するものを重ねたオーバーレイ画像を得るステップとを含む、コンピュータ読取可能な記録媒体。

【請求項4】 前記算出するステップは、

前記第1の画像フレームおよび前記第2の画像フレームの、第1のサブバンド画像から得られた前記オーバーレイ画像のテクスチャマップのクロス相関計算を行なうことにより、前記第2の画像フレームの前記第1のサブバンド画像に対応する第1の前記粗メッシュモデルの各ノードと、前記第1の画像フレームの前記第1のサブバンド画像に対応する前記第1の粗メッシュモデルの各ノードとを対応付け、前記第1の粗メッシュモデルの各ノードの前記第1の画像フレームから前記第2の画像フレームの間の相対的位置変化を算出するステップと、

前記第2の画像フレームの、前記第1の粗メッシュモデルよりも高いノード密度を有する第2の粗メッシュモデルに対して、前記第1の粗メッシュモデルの各ノードの位置変化に応じた変形を行なうステップと、前記第2の粗メッシュモデルは、前記第1のサブバンド画像よりも高い空間周波数を有する第2のサブバンド画像に対応するものであり、

前記第1の画像フレームおよび前記第2の画像フレームの、前記第2のサブバンド画像から得られた前記オーバーレイ画像のテクスチャマップのクロス相関計算を行なうことにより、前記第2の画像フレームの前記第1のサブバンド画像に対応する前記第2の粗メッシュモデルの各ノードと、前記第1の画像フレームの前記第2のサブバンド画像に対応する前記第2の粗メッシュモデルの各ノードとを対応付け、前記第2の粗メッシュモデルの各ノードの前記第1の画像フレームか

ら前記第 2 の画像フレームの間の相対的位置変化を算出するステップとを含む、請求項 3 に記載のコンピュータ読取可能な記録媒体。

【請求項 5】 先行する第 1 の画像フレームおよび前記第 1 の画像フレームより後の第 2 の画像フレームを含む、対象物の画像フレームのシーケンスから、前記対象物の表面の各部位の動きを測定するための表面動き測定装置であって、

相互の隣接関係が定められた複数個のノードを含むメッシュモデルによって前記第 1 の画像フレームの画像内の前記対象物の表面の各部位の位置を表現するための手段と、

予め定められた複数通りの解像度で前記第 1 の画像フレームの画像および前記第 2 の画像フレームの画像を各々複数個のサブバンド画像に変換するための手段と、

前記メッシュモデルから、前記メッシュモデルのノード密度より低く、かつ互いに異なるノード密度を有する複数個の粗メッシュモデルを生成するための手段と、前記複数個の粗メッシュモデルは、それぞれ前記複数個のサブバンド画像と対応付けられており、

前記第 1 の画像フレームおよび前記第 2 の画像フレームの各々の前記複数個のサブバンド画像に対してそれぞれ前記複数個の粗メッシュモデルのうちの対応するものを重ねたオーバレイ画像を得るための手段と、

前記第 1 の画像フレームおよび前記第 2 の画像フレームの、それぞれ対応するサブバンド画像から得られた前記オーバレイ画像のテクスチャマップのクロス相関を計算することにより、前記第 2 の画像フレームの前記メッシュモデルの各ノードと、前記第 1 の画像フレームの前記メッシュモデルの各ノードとを対応付け、各ノードの前記第 1 の画像フレームから前記第 2 の画像フレームの間の相対的位置変化を算出するための手段とを含む、表面動き測定装置。

【請求項 6】 前記算出するための手段は、

前記第 1 の画像フレームおよび前記第 2 の画像フレームの、第 1 のサブバンド画像から得られた前記オーバレイ画像のテクスチャマップのクロス相関計算を行なうことにより、前記第 2 の画像フレームの前記第 1 のサブバンド画像に対応する第 1 の前記粗メッシュモデルの各ノードと、前記第 1 の画像フレームの前記第

1 のサブバンド画像に対応する前記第 1 の粗メッシュモデルの各ノードとを対応付け、前記第 1 の粗メッシュモデルの各ノードの前記第 1 の画像フレームから前記第 2 の画像フレームの間の相対的位置変化を算出するための手段と、

前記第 2 の画像フレームの、前記第 1 の粗メッシュモデルよりも高いノード密度を有する第 2 の粗メッシュモデルに対して、前記第 1 の粗メッシュモデルの各ノードの位置変化に応じた変形を行なうための手段と、前記第 2 の粗メッシュモデルは、前記第 1 のサブバンド画像よりも高い空間周波数を有する第 2 のサブバンド画像に対応するものであり、

前記第 1 の画像フレームおよび前記第 2 の画像フレームの、前記第 2 のサブバンド画像から得られた前記オーバレイ画像のテクスチャマップのクロス相関計算を行なうことにより、前記第 2 の画像フレームの前記第 1 のサブバンド画像に対応する前記第 2 の粗メッシュモデルの各ノードと、前記第 1 の画像フレームの前記第 2 のサブバンド画像に対応する前記第 2 の粗メッシュモデルの各ノードとを対応付け、前記第 2 の粗メッシュモデルの各ノードの前記第 1 の画像フレームから前記第 2 の画像フレームの間の相対的位置変化を算出するための手段とを含む、請求項 5 に記載の表面動き測定装置。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

この発明は人物の表情を測定する処理に関し、特に、ビデオカメラ等から得られた画像中の人物の表情をリアルタイムで測定するための画像認識の分野に関する。

【0 0 0 2】

【従来の技術】

人間の間のコミュニケーション手段として、人物の顔の表情というのは非常に重要である。同様に、いわゆるマンーマシンインタフェースにおいても人物の表情は非常に大きな意味を持つ。

【0 0 0 3】

人物の表情は、人と人とはが相対面して会話するときには様々なレベルの情報を

伝達する。伝達される情報は、その人の感情の状態、会話における立場を含み、さらに発話中には、発声される言語音に関する音声上の特徴などに関する豊富な情報を含む。

【0004】

従来、ビデオ画像から人物の表情を分析するための試みが様々になされて来たが、ほとんどは定量的な評価を用いないものか、定量的な評価をしたとしても予め定められた非常に限定された顔の特徴に関する限定された定量的評価しか用いないものであった。もちろんそうした従来技術を用いても有用な結果を得ることはできるが、従来技術によればフレームごとに煩雑な検査をすることが必要であり、しかもたとえば人物の表情の動的属性を符号化することを考えた場合、十分に正確に表情の動きを測定することはできない。

【0005】

一方、人物の顔に直接張り付けたマーカを追跡することで、人物の顔の各部位の三次元的位置を正確に測定するための技術も存在している。しかしこうした技術は以下のような欠点を有する。

【0006】

(1) 測定のための機器が専用のものであって汎用性がなく、用途が限定されている。また、顔に直接マーカを張り付けるといふ、被験者にとっては不愉快な作業が必要なため、被験者の協力が不可欠である。さらに、この方法は設備が完備した場所、たとえば研究室内では使用できるが、一旦外部に出ればほとんど不可能である。

【0007】

(2) 人物の顔にマーカを取り付けることが可能な位置の数が限定されるので、どの箇所にマーカを取り付ければ最も効率良く正確な測定ができるかを知る必要がある。しかし現在のところ、どこにマーカを取り付けることが最適かについて知る方法はない。

【0008】

【発明が解決しようとする課題】

人物の表情を測定するもう一つのアプローチは、人物の顔のビデオ画像から表

情を定量的に表す情報を得ることである。ビデオ撮影は撮影機器さえあれば可能であり、人物の顔にマーカを付す必要もないので、上記した従来技術にまつわる問題を解消することが可能である。

【0 0 0 9】

しかし、ビデオ画像を用いたアプローチでは、撮影時に3次元情報を2次元情報に変換することによる画像の歪み、空間的解像度が低いこと、および照明状態による効果が非線形的であること、など多くの問題がある。また、こうした問題は人物の表情を測定する上で重要なことではあるが、人物以外の対象物の表面の動き、たとえば人物以外の動物の表情、物体の歪み、などの定量的な測定においても重要であり、これらにおいても同様の問題の解決が望まれる。

【0 0 1 0】

それゆえに本発明の目的は、ビデオ画像から、信頼性高くかつリアルタイムで人物などの対象物の表面の動きに関する定量的な情報を抽出することができる表面動き測定装置、そのための方法および当該方法をコンピュータを用いて実現するためのプログラムを記録した記録媒体を提供することである。

【0 0 1 1】

この発明の他の目的は、ビデオ画像から、信頼性高くかつリアルタイムで対象物の表面の各部位の移動に関する定量的な情報を抽出することができる表面動き測定装置、そのための方法および当該方法をコンピュータを用いて実現するためのプログラムを記録した記録媒体を提供することである。

【0 0 1 2】

この発明のさらに他の目的は、ビデオ画像から、対象物の表面のどの部位に関しても信頼性高くかつリアルタイムでその移動に関する定量的な情報を抽出することができる表面動き測定装置、そのための方法および当該方法をコンピュータを用いて実現するためのプログラムを記録した記録媒体を提供することである。

【0 0 1 3】

【課題を解決するための手段】

請求項1に記載の発明にかかる表面動き測定方法は、先行する第1の画像フレームおよび第1の画像フレームより後の第2の画像フレームを含む、対象物の画

像フレームのシーケンスから、対象物の表面の各部位の動きを測定するための表面動き測定方法であって、相互の隣接関係が定められた複数個のノードを含むメッシュモデルによって第1の画像フレームの画像内の対象物の表面の各部位の位置を表現するステップと、予め定められた複数通りの解像度で第1の画像フレームの画像および第2の画像フレームの画像を各々複数個のサブバンド画像に変換するステップと、メッシュモデルから、メッシュモデルのノード密度より低く、かつ互いに異なるノード密度を有する複数個の粗メッシュモデルを生成するステップと、複数個の粗メッシュモデルは、それぞれ複数個のサブバンド画像と対応付けられており、第1の画像フレームおよび第2の画像フレームの各々の複数個のサブバンド画像に対してそれぞれ複数個の粗メッシュモデルのうちの対応するものを重ねたオーバーレイ画像を得るステップと、第1の画像フレームおよび第2の画像フレームの、それぞれ対応するサブバンド画像から得られたオーバーレイ画像のテクスチャマップのクロス相関を計算することにより、第2の画像フレームのメッシュモデルの各ノードと、第1の画像フレームのメッシュモデルの各ノードとを対応付け、各ノードの第1の画像フレームから第2の画像フレームの間の相対的位置変化を算出するステップとを含む。

【0014】

請求項1に記載の発明によれば、サブバンド画像の異なるレベルごとに、対応する粗メッシュモデルを用いて第2の画像フレームの各ノードと、第1の画像フレームの各ノードとの対応関係が得られる。このとき、サブバンド画像と粗メッシュモデルのノード密度が対応しているために、顔の各部位について大きな動きは最も低いノード密度の粗メッシュモデルで位置を算出でき、より細かい動きについてはより高いノード密度の粗メッシュモデルで位置を算出できる。その結果、動きの大小にかかわらず画像フレームから対象物の表面部位の動きを算出することができる表面動き測定方法を提供できる。

【0015】

請求項2に記載の発明にかかる表面動き測定方法は、請求項1に記載の発明の構成に加えて、算出するステップは、第1の画像フレームおよび第2の画像フレームの、第1のサブバンド画像から得られたオーバーレイ画像のテクスチャマップ

のクロス相関計算を行なうことにより、第2の画像フレームの第1のサブバンド画像に対応する第1の粗メッシュモデルの各ノードと、第1の画像フレームの第1のサブバンド画像に対応する第1の粗メッシュモデルの各ノードとを対応付け、第1の粗メッシュモデルの各ノードの第1の画像フレームから第2の画像フレームの間の相対的位置変化を算出するステップと、第2の画像フレームの、第1の粗メッシュモデルよりも高いノード密度を有する第2の粗メッシュモデルに対して、第1の粗メッシュモデルの各ノードの位置変化に応じた変形を行なうステップと、第2の粗メッシュモデルは、第1のサブバンド画像よりも高い空間周波数を有する第2のサブバンド画像に対応するものであり、第1の画像フレームおよび第2の画像フレームの、第2のサブバンド画像から得られたオーバーレイ画像のテクスチャマップのクロス相関計算を行なうことにより、第2の画像フレームの第1のサブバンド画像に対応する第2の粗メッシュモデルの各ノードと、第1の画像フレームの第2のサブバンド画像に対応する第2の粗メッシュモデルの各ノードとを対応付け、第2の粗メッシュモデルの各ノードの第1の画像フレームから第2の画像フレームの間の相対的位置変化を算出するステップとを含む。

【0016】

請求項2に記載の発明によれば、請求項1に記載の発明の作用に加えて、低いノード密度の粗メッシュモデルを用いて得られた第2の画像フレームの画像の各部位の動きにより、より高い密度の粗メッシュモデルを用いた各部位の動きの算出時の初期の粗メッシュモデルの各ノードのおおまかな位置を推定し予め変形できる。そのため、この高い密度の粗メッシュモデルを用いて行なうクロス相関における計算量が減少し、高速に各部位の位置を算出することができる。

【0017】

請求項3に記載の発明にかかるコンピュータ読取可能な記録媒体は、先行する第1の画像フレームおよび第1の画像フレームより後の第2の画像フレームを含む、対象物の画像フレームのシーケンスから、対象物の表面の各部位の動きを測定するための表面動き測定方法を実現するようコンピュータを動作させるためのプログラムを記憶した、コンピュータ読取可能な記録媒体であって、表面動き測定方法は、相互の隣接関係が定められた複数個のノードを含むメッシュモデルに

よって第 1 の画像フレームの画像内の対象物の表面の各部位の位置を表現するステップと、予め定められた複数通りの解像度で第 1 の画像フレームの画像および第 2 の画像フレームの画像を各々複数個のサブバンド画像に変換するステップと、メッシュモデルから、メッシュモデルのノード密度より低く、かつ互いに異なるノード密度を有する複数個の粗メッシュモデルを生成するステップと、複数個の粗メッシュモデルは、それぞれ複数個のサブバンド画像と対応付けられており、第 1 の画像フレームおよび第 2 の画像フレームの各々の複数個のサブバンド画像に対してそれぞれ複数個の粗メッシュモデルのうちの対応するものを重ねたオーバーレイ画像を得るステップとを含む。

【 0 0 1 8 】

請求項 3 に記載の発明によれば、サブバンド画像の異なるレベルごとに、対応する粗メッシュモデルを用いて第 2 の画像フレームの各ノードと、第 1 の画像フレームの各ノードとの対応関係が得られる。このとき、サブバンド画像と粗メッシュモデルのノード密度が対応しているために、顔の各部位について大きな動きは最も低いノード密度の粗メッシュモデルで位置を算出でき、より細かい動きについてはより高いノード密度の粗メッシュモデルで位置を算出できる。その結果、動きの大小にかかわらず画像フレームから対象物の表面部位の動きを算出することができるコンピュータ読取可能な記録媒体を提供できる。

【 0 0 1 9 】

請求項 4 に記載の発明にかかるコンピュータ読取可能な記録媒体は、請求項 3 に記載の発明の構成に加えて、算出するステップは、第 1 の画像フレームおよび第 2 の画像フレームの、第 1 のサブバンド画像から得られたオーバーレイ画像のテクスチャマップのクロス相関計算を行なうことにより、第 2 の画像フレームの第 1 のサブバンド画像に対応する第 1 の粗メッシュモデルの各ノードと、第 1 の画像フレームの第 1 のサブバンド画像に対応する第 1 の粗メッシュモデルの各ノードとを対応付け、第 1 の粗メッシュモデルの各ノードの第 1 の画像フレームから第 2 の画像フレームの間の相対的位置変化を算出するステップと、第 2 の画像フレームの、第 1 の粗メッシュモデルよりも高いノード密度を有する第 2 の粗メッシュモデルに対して、第 1 の粗メッシュモデルの各ノードの位置変化に応じた変

形を行なうステップと、第 2 の粗メッシュモデルは、第 1 のサブバンド画像よりも高い空間周波数を有する第 2 のサブバンド画像に対応するものであり、第 1 の画像フレームおよび第 2 の画像フレームの、第 2 のサブバンド画像から得られたオーバレイ画像のテクスチャマップのクロス相関計算を行なうことにより、第 2 の画像フレームの第 1 のサブバンド画像に対応する第 2 の粗メッシュモデルの各ノードと、第 1 の画像フレームの第 2 のサブバンド画像に対応する第 2 の粗メッシュモデルの各ノードとを対応付け、第 2 の粗メッシュモデルの各ノードの第 1 の画像フレームから第 2 の画像フレームの間の相対的位置変化を算出するステップとを含む。

【 0 0 2 0 】

請求項 4 に記載の発明によれば、請求項 3 に記載の発明の作用に加えて、低いノード密度の粗メッシュモデルを用いて得られた第 2 の画像フレームの画像の各部位の動きにより、より高い密度の粗メッシュモデルを用いた各部位の動きの算出時の初期の粗メッシュモデルの各ノードのおおまかな位置を推定し予め変形できる。そのため、この高い密度の粗メッシュモデルを用いて行なうクロス相関における計算量が減少し、高速に各部位の位置を算出することができる。

【 0 0 2 1 】

請求項 5 に記載の発明にかかる表面動き測定装置は、先行する第 1 の画像フレームおよび第 1 の画像フレームより後の第 2 の画像フレームを含む、対象物の画像フレームのシーケンスから、対象物の表面の各部位の動きを測定するための表面動き測定装置であって、相互の隣接関係が定められた複数個のノードを含むメッシュモデルによって第 1 の画像フレームの画像内の対象物の表面の各部位の位置を表現するための手段と、予め定められた複数通りの解像度で第 1 の画像フレームの画像および第 2 の画像フレームの画像を各々複数個のサブバンド画像に変換するための手段と、メッシュモデルから、メッシュモデルのノード密度より低く、かつ互いに異なるノード密度を有する複数個の粗メッシュモデルを生成するための手段と、複数個の粗メッシュモデルは、それぞれ複数個のサブバンド画像と対応付けられており、第 1 の画像フレームおよび第 2 の画像フレームの各々の複数個のサブバンド画像に対してそれぞれ複数個の粗メッシュモデルのうちの対

応するものを重ねたオーバーレイ画像を得るための手段と、第1の画像フレームおよび第2の画像フレームの、それぞれ対応するサブバンド画像から得られたオーバーレイ画像のテクスチャマップのクロス相関を計算することにより、第2の画像フレームのメッシュモデルの各ノードと、第1の画像フレームのメッシュモデルの各ノードとを対応付け、各ノードの第1の画像フレームから第2の画像フレームの間の相対的位置変化を算出するための手段とを含む。

【0022】

請求項5に記載の発明によれば、サブバンド画像の異なるレベルごとに、対応する粗メッシュモデルを用いて第2の画像フレームの各ノードと、第1の画像フレームの各ノードとの対応関係が得られる。このとき、サブバンド画像と粗メッシュモデルのノード密度が対応しているために、顔の各部位について大きな動きは最も低いノード密度の粗メッシュモデルで位置を算出でき、より細かい動きについてはより高いノード密度の粗メッシュモデルで位置を算出できる。その結果、動きの大小にかかわらず画像フレームから対象物の表面部位の動きを算出することができる表面動き測定装置を提供できる。

【0023】

請求項6に記載の発明にかかる表面動き測定装置は、請求項5に記載の発明の構成に加えて、算出するための手段は、第1の画像フレームおよび第2の画像フレームの、第1のサブバンド画像から得られたオーバーレイ画像のテクスチャマップのクロス相関計算を行なうことにより、第2の画像フレームの第1のサブバンド画像に対応する第1の粗メッシュモデルの各ノードと、第1の画像フレームの第1のサブバンド画像に対応する第1の粗メッシュモデルの各ノードとを対応付け、第1の粗メッシュモデルの各ノードの第1の画像フレームから第2の画像フレームの間の相対的位置変化を算出するための手段と、第2の画像フレームの、第1の粗メッシュモデルよりも高いノード密度を有する第2の粗メッシュモデルに対して、第1の粗メッシュモデルの各ノードの位置変化に応じた変形を行なうための手段と、第2の粗メッシュモデルは、第1のサブバンド画像よりも高い空間周波数を有する第2のサブバンド画像に対応するものであり、第1の画像フレームおよび第2の画像フレームの、第2のサブバンド画像から得られたオーバー

イ画像のテクスチャマップのクロス相関計算を行なうことにより、第 2 の画像フレームの第 1 のサブバンド画像に対応する第 2 の粗メッシュモデルの各ノードと、第 1 の画像フレームの第 2 のサブバンド画像に対応する第 2 の粗メッシュモデルの各ノードとを対応付け、第 2 の粗メッシュモデルの各ノードの第 1 の画像フレームから第 2 の画像フレームの間の相対的位置変化を算出するための手段とを含む。

【0 0 2 4】

請求項 6 に記載の発明によれば、請求項 5 に記載の発明の作用に加えて、低いノード密度の粗メッシュモデルを用いて得られた第 2 の画像フレームの画像の各部位の動きにより、より高い密度の粗メッシュモデルを用いた各部位の動きの算出時の初期の粗メッシュモデルの各ノードのおおまかな位置を推定し予め変形できる。そのため、この高い密度の粗メッシュモデルを用いて行なうクロス相関における計算量が減少し、高速に各部位の位置を算出することができる。

【0 0 2 5】

【発明の実施の形態】

以下、本発明の実施の形態 1 にかかる表情測定装置について説明する。この表情測定装置は、パーソナルコンピュータまたはワークステーション等、コンピュータ上で実行されるソフトウェアにより実現されるものであって、人物の顔のビデオ画像から、人物の顔の各部の動き（表情）を測定するためのものである。図 1 に、この表情測定装置の外観を示す。

【0 0 2 6】

図 1 を参照してこのシステム 2 0 は、CD-ROM (Compact Disc Read-Only Memory) ドライブ 5 0 および FD (Flexible Disk) ドライブ 5 2 を備えたコンピュータ本体 4 0 と、コンピュータ本体 4 0 に接続された表示装置としてのディスプレイ 4 2 と、同じくコンピュータ本体 4 0 に接続された入力装置としてのキーボード 4 6 およびマウス 4 8 と、コンピュータ本体 4 0 に接続された、画像を取込むためのビデオカメラ 3 0 とを含む。この実施の形態の装置では、ビデオカメラ 3 0 としては CCD (固体撮像素子) を含むビデオカメラを用い、ビデオカメラ 3 0 の前の人物の表情を測定する処理を行うものとする。

【 0 0 2 7 】

図 2 に、このシステム 2 0 の構成をブロック図形式で示す。図 2 に示されるようにこのシステム 2 0 を構成するコンピュータ本体 4 0 は、CD-ROM ドライブ 5 0 および FD ドライブ 5 2 に加えて、それぞれバス 6 6 に接続された CPU (Central Processing Unit) 5 6 と、ROM (Read Only Memory) 5 8 と、RAM (Random Access Memory) 6 0 と、ハードディスク 5 4 と、ビデオカメラ 3 0 からの画像を取込むための画像取込装置 6 8 とを含んでいる。CD-ROM ドライブ 5 0 には CD-ROM 6 2 が装着される。FD ドライブ 5 2 には FD 6 4 が装着される。

【 0 0 2 8 】

既に述べたようにこの表情測定装置の主要部は、コンピュータハードウェアと、CPU 5 6 により実行されるソフトウェアとにより実現される。一般的にこうしたソフトウェアは CD-ROM 6 2、FD 6 4 等の記憶媒体に格納されて流通し、CD-ROM ドライブ 5 0 または FD ドライブ 5 2 等により記憶媒体から読取られてハードディスク 5 4 に一旦格納される。または、当該装置がネットワークに接続されている場合には、ネットワーク上のサーバから一旦ハードディスク 5 4 に一旦コピーされる。そうしてさらにハードディスク 5 4 から RAM 6 0 に読出されて CPU 5 6 により実行される。なお、ネットワーク接続されている場合には、ハードディスク 5 4 に格納することなく RAM 6 0 に直接ロードして実行するようにしてもよい。

【 0 0 2 9 】

図 1 および図 2 に示したコンピュータのハードウェア自体およびその動作原理は一般的なものである。したがって、本発明の最も本質的な部分は FD ドライブ 5 2、FD 6 4、ハードディスク 5 4 等の記憶媒体に記憶されたソフトウェアである。

【 0 0 3 0 】

なお、最近の一般的傾向として、コンピュータのオペレーティングシステムの一部として様々なプログラムモジュールを用意しておき、アプリケーションプログラムはこれらモジュールを所定の配列で必要な時に呼び出して処理を進める方

式が一般的である。そうした場合、当該表情測定装置を実現するためのソフトウェア自体にはそうしたモジュールは含まれず、当該コンピュータでオペレーティングシステムと協働してはじめて表情測定装置が実現することになる。しかし、一般的なプラットフォームを使用する限り、そうしたモジュールを含ませたソフトウェアを流通させる必要はなく、それらモジュールを含まないソフトウェア自体およびそれらソフトウェアを記録した記録媒体（およびそれらソフトウェアがネットワーク上を流通する場合のデータ信号）が実施の形態を構成すると考えることができる。

【0031】

上記したように、本願発明は実質的にはソフトウェア処理によって実現される。本実施の形態の装置において人物の顔の動き（表情）を測定する処理を行なうソフトウェアは以下のような制御構造を有する。

【0032】

図3を参照して、まずモーショントラッキングのための種々の準備を実行する初期処理を行なう（80）。初期処理の内容に関しては図4を参照して後述するが、特にここでは後の人物の顔の各部位の動きを測定するための、人物の顔の基準画像が決定される。またこの初期処理では、基準画像の顔領域にフィットする、メッシュが付与される楕円が決定される。初期処理が完了した後、ビデオ画像の1フレームを取り込んでフレームごとの処理（モーショントラッキング）を行ない、基準フレームに対する現フレームにおける人物の顔の各部位を示す情報を出力する（82）。続いてビデオ画像の全フレームが終了したか否かを判定する（84）。フレームの入力がこれ以後なければ処理を終了し、次のフレームがあれば再びステップ82以下の処理を繰返す。

【0033】

こうして、ステップ82によって基準フレームに対する現フレームの顔画像の各部位の動きを表す情報が時系列的に得られる。入力されるフレームごとに結果が得られるので、このシステムはリアルタイムにビデオ画像から人物の表情を測定する処理に適している。

【0034】

図4を参照して、図3のステップ80で行なわれる初期処理は以下の手順で実現される。まず、図1に示すビデオカメラ30からのビデオ画像の1フレームが読込まれる(90)。ここでは、予め撮影された画像に対して処理を行なうものとし、ステップ90ではこのうちの最初のいくつかのフレームが読込まれるものとする。なお、ここで入力されるビデオ画像はどのようなフォーマットのものでよく、またカラー画像でもグレースケール画像でも同様の効果を得ることができる。

【0035】

ステップ92では、これら読込まれたフレームの中からどのフレームの画像を基準画像とするかが決定される。このフレームの選択は、通常のコンピュータにおけるユーザインタラクションを用いてオペレータにより行なわれる。ここで選択されたフレームを以下「基準フレーム」と呼ぶ。基準フレームは、この後の分析による顔の各部位の位置に対する基準となる。なお、この実施の形態では基準フレームをオペレータが選択することとしているが、何らかの基準を設けることによって基準フレームを自動的に決定するようにしてもよい。

【0036】

次に、基準フレームの顔画像に対して、その輪郭を決定する処理が行なわれる(94)。本実施の形態の装置では輪郭はオペレータによって、少なくとも画像のうち4点を指定することによって決定される。5点を指定した様子を図9に示す。図9において、指定された5点は周囲に黒い枠を付けた白抜きの「+」印で表されている。これら5点のうち2点は耳たぶ付近、他の2点は画面の上端近く、最後の1点はあご先である。

【0037】

これらの点が指定されると、所定のアルゴリズムによってこれらの点に最もよくフィットする楕円の中心が決定される(96)。この処理にはDownhill simplex minimization method(目標値との差を、制御対象の属性(位置、大きさ、その他の属性一般)を決定するパラメータの関数として表し、このずれが最小となるようにパラメータを定める方法)を用いる。決定された楕円の例を図10に示す。

【 0 0 3 8 】

この楕円は、その短軸の方向が両目の目尻を結ぶ線の方角と一致するように整列される。このためにステップ 9 8 で目尻をマーキングする処理がオペレータにより行なわれる。このようにして楕円の両軸を決定することにより、後述するメッシュモデルの方角が決定される。またこのようにして顔領域を決定する場合、被験者が変わっても一貫した結果を得ることができる。

【 0 0 3 9 】

こうして初期の楕円の両軸を決定したのち、再度ステップ 9 6 で行なった処理と類似の最適化法によって楕円の角度、位置、大きさなどを最適化するための微調整が行なわれる（1 0 0）。この処理によって、人物の顔領域とよく一致するようにその大きさと方角とが決定された楕円が得られる。

【 0 0 4 0 】

この段階でのアプローチには 2 つの主な特徴がある。第 1 に、最終的に得られる楕円に対する、顔の輪郭点のうちの特異的な点による影響が最低におさえられる。第 2 に、顔領域の一部が画面の外にはみ出しているも、顔領域を正確に指定することができる。

【 0 0 4 1 】

こうして顔領域にフィットする楕円が決定されると、予め設定可能な一連の初期値と、基準画像の顔領域から決定された楕円を特定するパラメータとを用いて、3 次元のメッシュモデルの初期値が計算される。3 次元メッシュモデルの例を図 1 1 に示す。このメッシュモデルは、3 次元の回転楕円体の表面に分布するノードの集合として定義される。各ノードは、楕円の表面上に所定間隔でメッシュをなして配置されるものとする。このメッシュモデルを顔画像に重畳して示したのが図 1 2 である。このメッシュモデルの各ノードの位置によって、人物の顔の各部位の 3 次元位置が表される。

【 0 0 4 2 】

この場合、最も留意すべきパラメータは、各メッシュノードの間に平均していくつの画素を配置するか、ということである。この数によって、最終的に測定される顔の各部位の位置の解像度が決定され、また後述するウェーブレット変換に

よって分析が行なわれる空間周波数の範囲が決定されるためである。

【0 0 4 3】

以上によって、顔の各部位の動きを測定するための基本的準備は完了である。

ただし、撮影に用いられるビデオ画像の光学的パラメータ（焦点距離など）によって画像には歪みが生じうるので、それらを補償するためにビデオカメラの射影幾何学的データを読込む（1 0 2）。また、人物がその表情を変化させる際には、頭の動きを伴うのが通常である。頭が動くと、必然的に顔領域も動くので、顔の各部位の動きを正確に測定するためには頭の動きを測定し頭の動きによる顔領域の動きを補償しなければならない。そのための頭の基準位置をステップ 1 0 4 で読込む。なお頭の位置を測定するために、以下で説明する実施例では頭部にそのためのセンサを装着している。しかしこの発明はそれには限定されず、ビデオ画像から直接に頭部の動きを検出するようにしてもよい。

【0 0 4 4】

さて、次に図 3 のステップ 8 2 で行なわれる各フレームごとのモーショントラッキング処理について説明する。このモーショントラッキング処理は、基本的にはフレーム間の顔面の小部分の位置の変化をそれぞれ決定することにより 1 フレームずつ行なわれる。図 5 を参照して、まず処理対象のフレーム（これを以下の説明では「フレーム n」とする。）を読込む（1 1 0）。

【0 0 4 5】

ステップ 1 1 2 で、読込まれたフレーム n の画像を構成する全画素に対して、多重解像度の 2 次元離散ウェーブレット変換（DWT）を施す（1 1 2）。この処理はちょうど、図 1 3 に示されるようにローパスフィルタとハイパスフィルタとの対からなる多階層の、かつその特性が分析の目的にしたがって調整できるフィルタバンクをフレーム n の画像を通していく処理に相当する。この装置では、コンパクト台を持つスプラインウェーブレットによる二重直交方式に対応するフィルタを使用する。

【0 0 4 6】

二次元 DWT 方式は 4 つの出力を持つ。第 1 は、信号の低周波数成分を含むいわゆる「近似」出力であって、これは次に一レベル下のウェーブレット変換に対

する入力となっている。第2～第4は、いわゆる「詳細」または「サブバンド」出力であって、これらは高周波数成分を含む。サブバンド出力は、対角方向、水平方向、および垂直方向の3つである。もとの画像イメージ（グレースケール）と、各レベルの近似およびサブバンド出力の関係を図13に示す。サブバンド出力は、各レベルにおけるハイパスフィルタの出力である。

【0047】

近似出力を得るためには、入力信号の画像のマトリクスを、行方向および列方向の双方に関してローパスフィルタに通す。対角方向サブバンド出力を得るためには、同じプロセスをハイパスフィルタを用いて行なう。垂直方向サブバンド出力を得るためには、入力信号の画像のマトリクスを、行に沿ってハイパスフィルタに、列に沿ってローパスフィルタに、それぞれ通す。水平方向サブバンド出力を得るためには、垂直方向サブバンド出力のときのフィルタを入れ替える。本実施の形態では、DWTの各レベルで3つのサブバンド出力のみを用いる。

【0048】

再び図5を参照して、続いてフレーム n に対するメッシュモデルをフレーム $n-1$ に対して得られたメッシュモデルから作成する（114）。次に、ステップ114で得られたメッシュモデルと、前フレームに対して最終的に得られたメッシュモデルとを、フレーム n に対して得られた頭の動きを補正するようにして画像平面に射影する（116）。さらにステップ118とステップ120とによって、ウェーブレット変換の各レベルに対して図6に示す後述処理を繰返して行なう。

【0049】

ステップ118の処理をウェーブレット変換の全レベルに対して行なうと、画像平面上でフレーム n のどのノードがフレーム $n-1$ のどのノードに対応するかを表す情報が得られている。この情報から、フレーム n において、顔領域の各部位が基準フレームの対応する部位に対してどれくらい移動しているかを判定することができる。この情報は二次元であるから、これをもとの顔領域に対応する3次元の値に逆射影する（122）。これにより、フレーム n での顔の各部位の3次元的位置が確定する。

【0 0 5 0】

続いて、次のフレームに対するステップ 1 1 6 の処理で使用するため、こうして求められたフレーム n のメッシュモデルを保存し、フレーム n のメッシュモデルの各ノードの位置データを出力して (1 2 6) このフレーム n に対するモーショントラッキング処理を終了する。

【0 0 5 1】

次に、図 6 を参照して、図 5 のステップ 1 1 8 で行なわれるウェーブレット変換の各レベルごとの処理について説明する。まず、メッシュモデルのノード数をウェーブレット変換のレベルにあわせて少なくする (1 3 0)。このようにノードの密度が低くされたメッシュモデルをこの明細書では「粗メッシュモデル」と呼ぶこととする。本実施の形態では各レベルで用いられるノードの密度は、レベルが 1 下がると 1 / 2 倍となるように選ばれている。処理はレベル 1 から始められる。なお図 1 3 に示すウェーブレット変換では、レベル 5 が最も高いレベルであり、レベル数が小さくなるにしたがってレベルは下がるものとする。レベルが上がるほど、ウェーブレット変換後の画像の空間周波数は高い。ただし以下の説明では簡略にするためレベルが 3 つの場合を想定する。

【0 0 5 2】

ここでの基本的考え方は以下のとおりである。ノード密度の低い粗メッシュモデルを用いて、フレーム n のウェーブレット変換後の画像の各ノードと、フレーム n - 1 のウェーブレット変換後の画像の各ノードとの間の相関を計算することにより、2 画像間での各ノードの対応関係が分かる。ノード密度が低いので、この処理により顔の各部位の比較的大きな動きが検出できる。ここでは、大きな動きは顔の大きい領域に影響を及ぼすということを仮定している。

【0 0 5 3】

次に、上記したように検出されたノード密度の低い粗メッシュモデルの各ノードの動きに基づいて、よりノード密度の高い粗メッシュモデルの各ノードの概略の位置が判定できる。特に、ノード密度の低い粗メッシュモデルには存在していないがノード密度の高い粗メッシュモデルには存在しているノードについては、ノード密度の低い粗メッシュモデルの各ノードの位置を補間することにより、そ

の位置が推定できる。こうして得られたよりノード密度の高い粗メッシュモデルを、対応するレベルのウェーブレット変換画像に重ねて再度フレーム $n-1$ の対応するレベルのウェーブレット変換画像および粗メッシュモデルとのクロス相関を計算することにより、このレベルでの各ノードの前フレーム $n-1$ からの動きが検出できる。

【0054】

こうした処理を最も高いレベルから最も低いレベルまで順に行なうことにより、最終的に顔の各部位の位置を細かい単位で検出することができる。

【0055】

この処理を実現するために、この「ウェーブレット変換の各レベルごとの処理」ではレベルの高いものから順にレベルを下げて処理を行なうようにし、かつ、前のレベルで得られたメッシュ密度の低い粗メッシュモデルを次のレベルのスタート時の粗メッシュモデルの作成に使用するようになっている。

【0056】

再び図6を参照して、ステップ130で処理対象のレベルに応じてメッシュノードの濃度を下げたメッシュノードに対して次に述べる処理を行なう。

【0057】

ステップ132で、フレーム n の画像に対するこのレベルのウェーブレット変換のサブバンド画像にこの粗メッシュモデルを重ね、フレーム $n-1$ のサブバンド画像に対するこのレベルのウェーブレット変換画像にフレーム $n-1$ の粗メッシュモデルを重ねてオーバーレイ画像を得る。

【0058】

次に、ステップ134およびステップ136によって、粗メッシュモデルに含まれる各ノードごとに後述する処理を行なってノード間の対応関係を定める。

【0059】

全てのノードについて上記した処理をおこなって、フレーム n の各ノードの位置を定められた対応関係にしたがって移動させることにより粗メッシュモデルを変形させる(138)。

【0060】

こうして得られた変形後の粗メッシュモデルについて、中間ノードの位置を補間によって計算し中間ノードを挿入する（140）。こうして得られた粗メッシュモデルが次のレベルでの処理に用いられる。これによりこの処理は終了する。

【0061】

図6のステップ134で各メッシュノードごとに行なわれる処理は図7のとおりである。

【0062】

まず、探索セグメントを決定する（150）。動き検出は、二つの連続する画像の顔のテクスチャマップを用いた2次元的相关計算により行なわれるので、そのために各メッシュノードの周辺の領域を適切に規定することが必要である。この領域を探索セグメントと呼ぶ。この実施の形態では、図8の上に示すように、フレームnのメッシュモデルのあるノード170に関連する探索セグメント160は、ノード170に隣接する4つのノード172、174、176および178により囲まれるポリゴン180である。

【0063】

サブサンプリングしないDWTによって、画像の各画素に対してある離散値（画素値）が計算される。上記したポリゴン180内の全画素の画素値によって、探索セグメント160のテクスチャマップが決定される。

【0064】

こうして決定された探索セグメントには、次に隣接するフレーム内の対応する探索セグメントに対して既に知られている変形にあわせるよう適合する処理が施される（152）。すなわち、図8の上に示した探索セグメント160が、隣接するフレームの対応する探索セグメント162と同じように変形される。この場合の適合処理は、各ポリゴン180を構成する要素182を単位として行なわれる。

【0065】

この変形は、隣接するノードがそれぞれ異なる動きをすることにより生ずると考えられる。この要因としては二つ考えられる。第1は頭部の動きによるものである。頭部の動きは画像面と平行とはいえないことが多いので、各ノードの動き

にずれが生じる。第2は前のウェーブレットレベルにおける、より粗な粗メッシュモデルによるモーショントラッキングの結果補間されたノード座標によって生じるものである。大きな変形は主としてこの第2の要因によって生じるものと考えられ、本発明においては口をあけたときなどの、大きな変形に対処するためにステップ140で採用されたものである。この変形（およびそれに伴う適合化処理）を考慮しないと、口をあけた場合の口の内部などには信頼のおけるテクスチャマップが存在しないために大きな誤差を生じるおそれがある。

【0066】

続いて、ステップ154で各サブバンドごとにフレーム n とフレーム $n-1$ との各ノードの探索セグメントに対応するテクスチャマップ間の相関計算が行なわれ、全サブバンドについて合計される。この相関計算によりノードの対応関係が得られ、その結果実際の顔の各部位（ノード）の位置の測定が行なわれたことになる。より具体的には、上記したクロス相関の結果、フレーム n のある探索セグメントとフレーム $n-1$ のある探索セグメントとが対応付けられ、フレーム n のその探索セグメントの中心ノードについてその移動量が決定される。これがそのノードの移動ベクトルとなる。

【0067】

なおここで、フレーム $n-1$ において得られた（変形された）粗メッシュモデルをフレーム n の画像に対して最初に適用することにより、顔の各部位の位置はおおまかながら既に推定されていることに注意する必要がある。すなわち、最も粗いレベルのウェーブレット変換によって頭部の動きによる影響が計算に入れられ、より密なレベルのウェーブレット変換では既にそれまでに求められていた動きの影響が計算に入れられている。

【0068】

たとえば、あご大きく開いたり閉じたりした結果、あご先全体が動いたとして、その動きは既に最も粗いレベルでは考慮され位置合わせが済んでいることになり、そのためたとえばあご先に近い下唇の一部の探索の開始位置はあご先の動きにあわせてずらされることになる。

【0069】

さらに、各部位の位置がほぼ予め推定できるだけでなく、探索セグメントの形状も予測でき、あわせてそのテクスチャマップがどのようなものとなっているかもほぼ予測が可能である。再度あごの大きな動きを例にとると、ちょうど口角部分に存在する比較的小さな探索セグメントの場合、そのテクスチャマップは非常に大きな変化を被るであろう。上述したより粗いレベルでのモーショントラッキングおよび適合処理の結果を用いることにより、この変化をほぼ予測することが可能となる。

【0070】

上述のように、探索セグメントの位置を推定するために、隣接する2つのフレームのサブバンド画像のテクスチャマップに対するクロス相関を用いている。このとき、前述のとおり各探索セグメントのおおまかな位置は判明しているので、フレーム n のある探索セグメントと、フレーム $n-1$ の全ての探索セグメントとの相関を計算する必要はない。予め判明しているおおまかな位置の周辺のみでよい。したがって密なレベルでの相関計算の計算量が減少し、精度の高さと比較して高速な判定を行なうことができる。計算による誤判定が生ずるおそれも少なくなる。したがって、問題は、いかにしてクロス相関の計算において、十分に精度高く、かつ計算コストも低くなるような制約条件が課された動きベクトルをどのようにして見い出するか、ということのみになる。

【0071】

この問題に対する一つの解決は、探索ノード内の各画素の位置を、探索ノードの中心に存在するノードを原点とした相対座標で表すことである。これによって、たとえば隣接するノードがこの中心ノードから離れている場合には動きベクトルによって検出できる動きの大きさは大きくなり、探索ノードが全体として圧縮されて隣接ノードが中心ノードに近い場合には小さなシフト量しか検出できないことになる。加えて、中心ノードに対して一方で隣接するノードが中心ノードから大きく離れ、他方で中心ノードに隣接するノードが中心ノードからそれほど離れていない場合にも同じ議論が成り立つ。

【0072】

この方法は、ビデオ画像の顔領域は連続性を有しているという基本的性質を用

いている。ほぼ同じ間隔だけ間をおいたメッシュから分析を開始するので、一旦大きな顔の動きにより大きく移動した部分も、やがてほとんどもとの位置に戻ってくる可能性が高いということを仮定することができる。またその逆も仮定できる。

【0073】

以上の処理によって、入力される一連のビデオ画像から、人物の顔の各部位の動きをリアルタイムで計算することができる。図14を用いて、再度上記した処理の全体的な流れを概観しておく。

【0074】

まずフレーム $n-1$ のメッシュモデル200を、フレーム $n-1$ 頭部の位置データなどを用いて画像平面に射影する(210)。この射影210に対し、最も粗いメッシュレベルの粗メッシュモデルを作成し(212)、それを最も低いレベルのウェーブレット変換のサブバンド画像に重ねてオーバーレイ画像を得る(214)。同様の処理を2番目に粗い粗メッシュモデル(222)および2番目のレベルのウェーブレット変換のサブバンド画像(224)、ならびに最も密な粗メッシュレベルのメッシュモデル(232)および3番目のレベルのウェーブレット変換のサブバンド画像に重ねる(234)。ここまではフレーム $n-1$ のデータを用いて行える処理である。

【0075】

フレーム n に対しては、まずフレーム $n-1$ のメッシュモデル200を、フレーム n の頭部の位置座標データなどを用いて座標平面に射影する(240)。このメッシュモデル240から、最も粗い粗メッシュモデル(242)を作成し、それを最も空間周波数の低いウェーブレット変換のサブバンド画像に重ねてオーバーレイ画像を得る(244)。このサブバンド画像244とフレーム $n-1$ のサブバンド画像214とのテクスチャマップを用い、各探索セグメントの間でクロス相関をとる(250)ことにより、フレーム n の最も粗いレベルでの各ノードの位置が推定される。

【0076】

次に、こうして推定された最も粗いレベルでのノードの移動を考慮して変形さ

れたメッシュモデル260を作成し、このメッシュモデル260のノード密度を粗メッシュモデル242の2倍まで落として粗メッシュモデル262を作成する。この粗メッシュモデルを2番目のレベルのウェーブレット変換のサブバンド画像と重ね(264)、さらにフレーム $n-1$ のサブバンド画像224との間でテクスチャマップを用いたクロス相関計算をする(270)ことで各探索セグメントのクロス計算を行なう。これにより、フレーム n の2番目に粗いレベルでの各ノードの位置が推定される。

【0077】

次に、こうして推定された2番目に粗いレベルでのノードの移動を考慮して変形されたメッシュモデル280を作成し、このメッシュモデル260のノード密度を粗メッシュモデル262の2倍まで落とした粗メッシュモデル282を作成する。この説明ではこの粗メッシュモデル282が最も高いノード密度を有する。この粗メッシュモデルを最も密なレベルの(最も空間周波数の高い)ウェーブレット変換のサブバンド画像と重ね(284)、さらにフレーム $n-1$ のサブバンド画像234との間でテクスチャマップを用いた相関計算をすることで各探索セグメントのクロス相関計算を行なう(290)。これにより、フレーム n の最も密なレベルでの各ノードの位置が推定される。こうして得られた各探索ノードの移動ベクトルを用いてもとのメッシュモデル240を変形し(300)、3次元に逆射影することによりフレーム n のメッシュモデルが得られる(302)。

【0078】

なお、以上の実施の形態は人間の顔の各部位の動きをビデオ画像から測定する機能を有するものである。しかしこの発明はこれには限定されず、人間の顔と同じように部分的に細かい動きを生ずるような物体、動物などの表面の動きを測定する装置および方法一般に適用することができる。

【0079】

また、上の説明ではフレームごとに上記した処理を行なうこととしたが、飛越し走査をする画像信号の場合には、フィールドごとに上記した処理を行なうようにしてもよい。

【0080】

今回開示された実施の形態はすべての点で例示であって制限的なものではないと考えられるべきである。本発明の範囲は上記した説明ではなくて特許請求の範囲によって示され、特許請求の範囲と均等の意味および範囲内でのすべての変更が含まれることが意図される。

【図面の簡単な説明】

【図 1】 本発明の実施の形態 1 にかかるシステムの外觀図である。

【図 2】 本発明の実施の形態 1 にかかるシステムのハードウェア的構成を示すブロック図である。

【図 3】 本発明の実施の形態 1 にかかるシステムで実行される処理の全体のフローチャートである。

【図 4】 本発明の実施の形態 1 にかかるシステムで実行される初期処理のフローチャートである。

【図 5】 本発明の実施の形態 1 にかかるシステムで実行されるモーショントラッキング処理のフローチャートである。

【図 6】 図 5 に示されるウェーブレット変換の各レベルごとの処理をさらに詳細に示すフローチャートである。

【図 7】 図 6 に示される各メッシュノードごとの処理をさらに詳細に示すフローチャートである。

【図 8】 探索セグメントの適合処理を説明するための図である。

【図 9】 初期処理において人物の目尻を指定する処理を示す図である。

【図 10】 初期処理において、指定された人物の両目尻から顔領域にフィットする楕円を推定し描いた状態を示す図である。

【図 11】 人物の顔にフィットされる回転楕円体と回転楕円体に付与されるメッシュとを三次元的に示す図である。

【図 12】 人物の顔画像にメッシュが付与された回転楕円体をフィットさせた状態を示す図である。

【図 13】 ウェーブレット変換の各レベルで得られる情報を示す図である。

【図 14】 本願発明の実施の形態のシステムにおける情報間の関係を示す

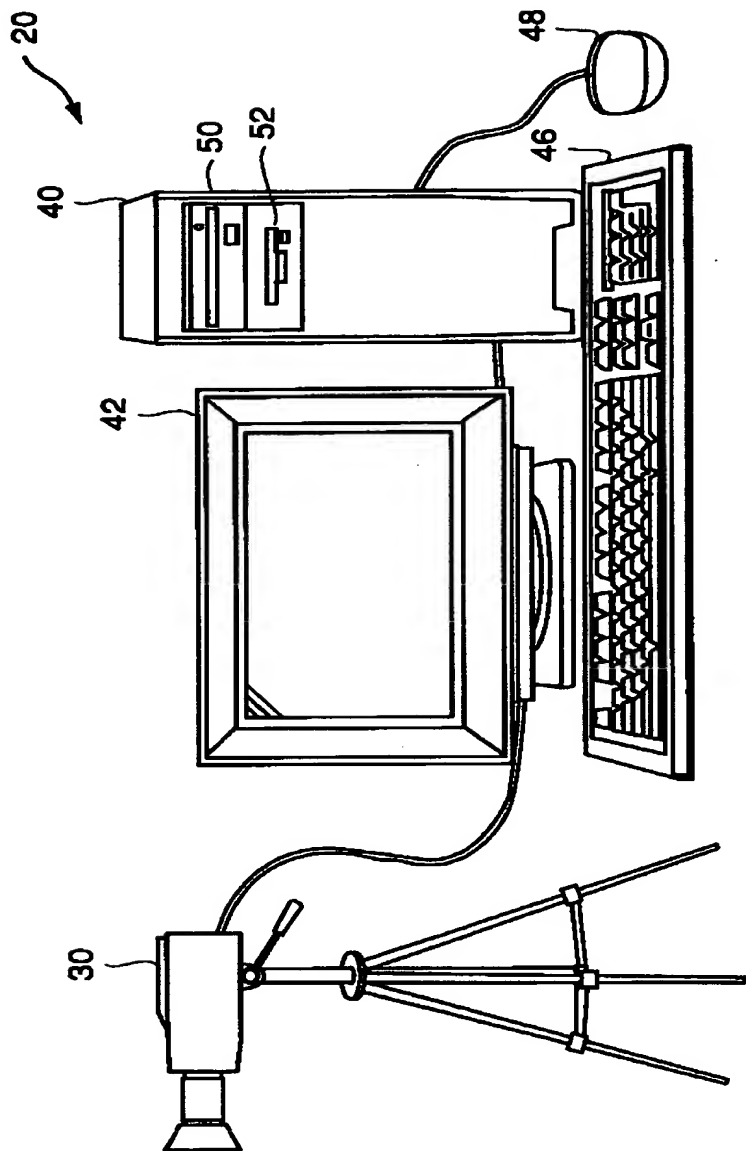
図である。

【符号の説明】

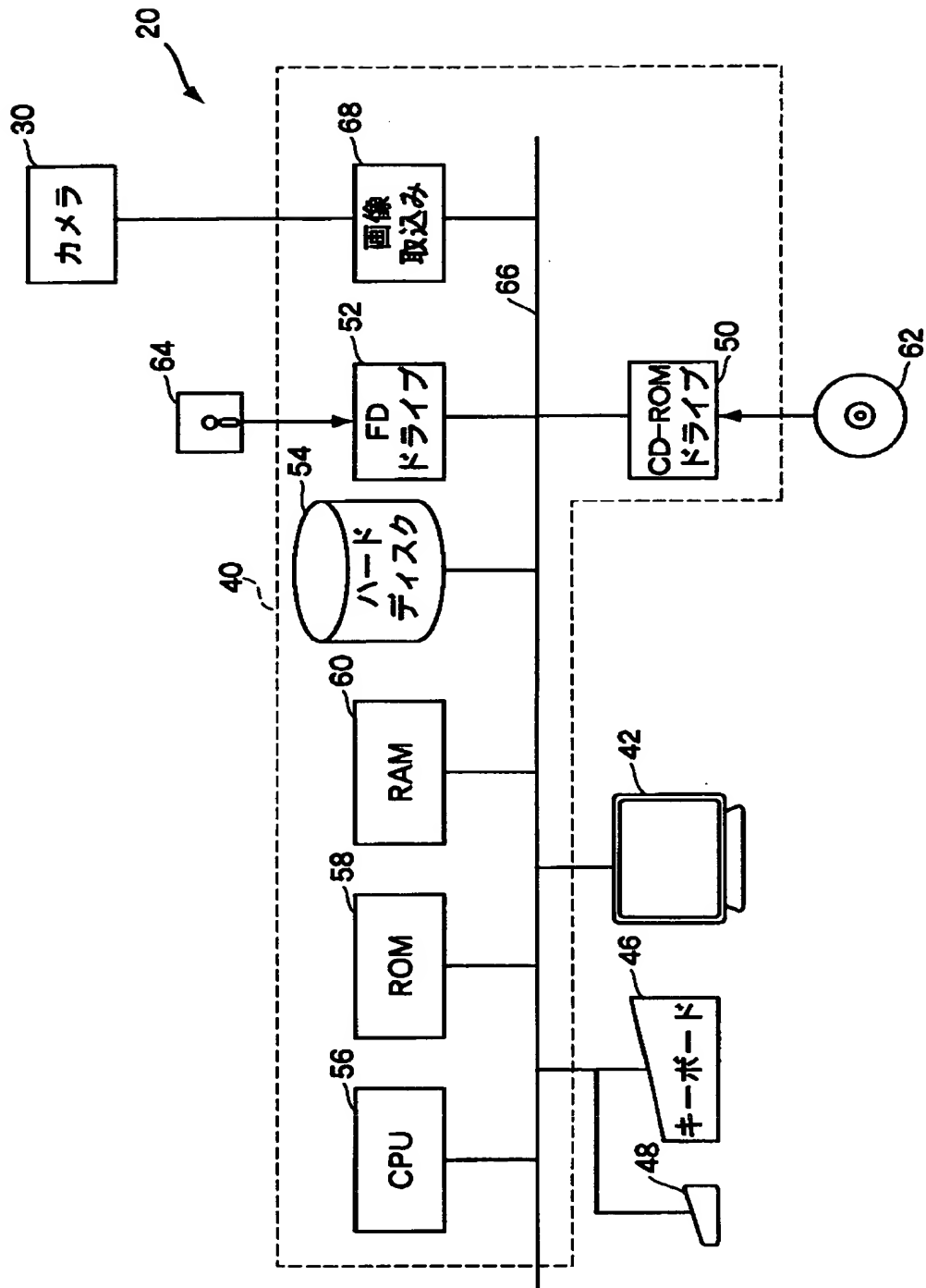
20 表情測定装置、30 カメラ、40 コンピュータ本体、42 モニタ、160、162 探索領域、170、172、174、176、178 ノード、214、224、234、244、264、284 サブバンド画像、212、222、232、242、262、282 粗メッシュモデル。

【書類名】 図面

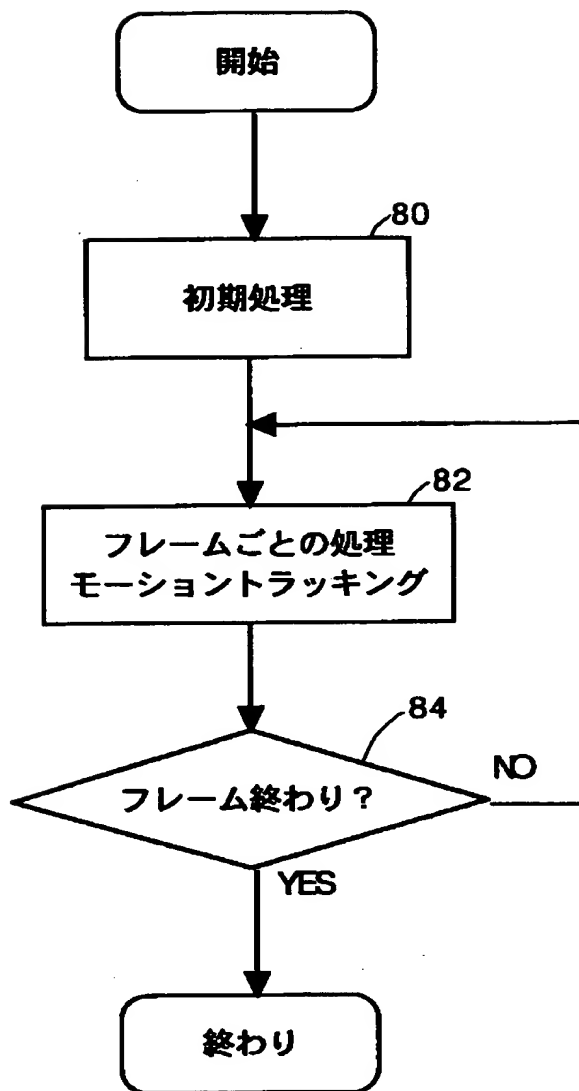
【図 1】



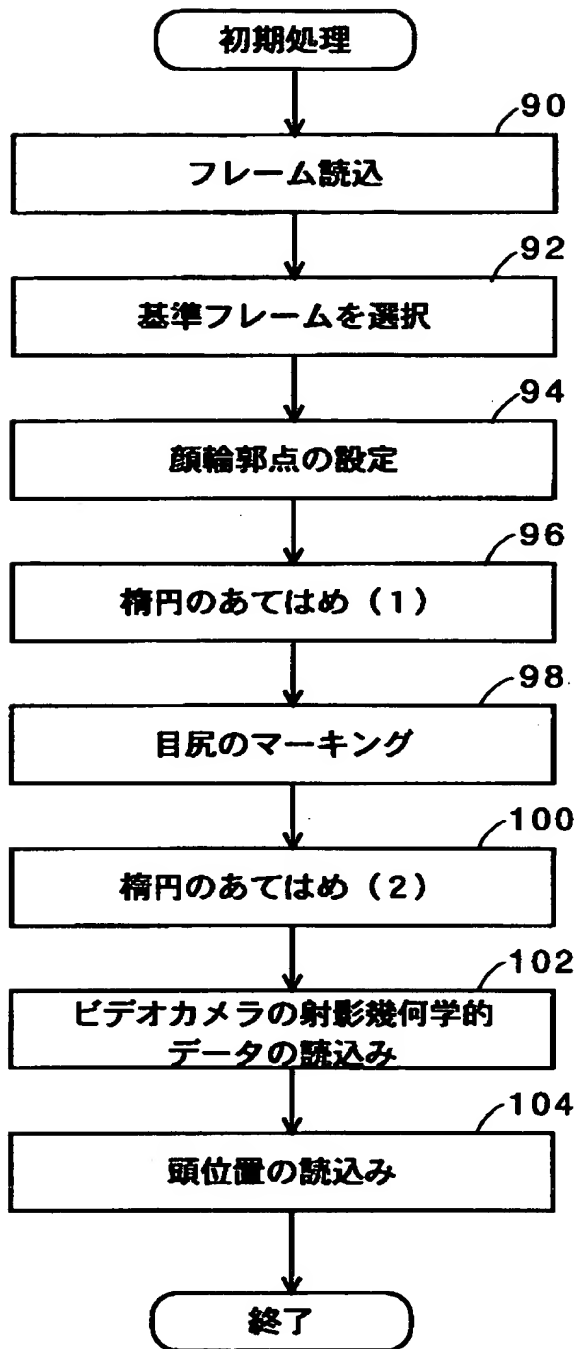
【図 2】



【図 3】

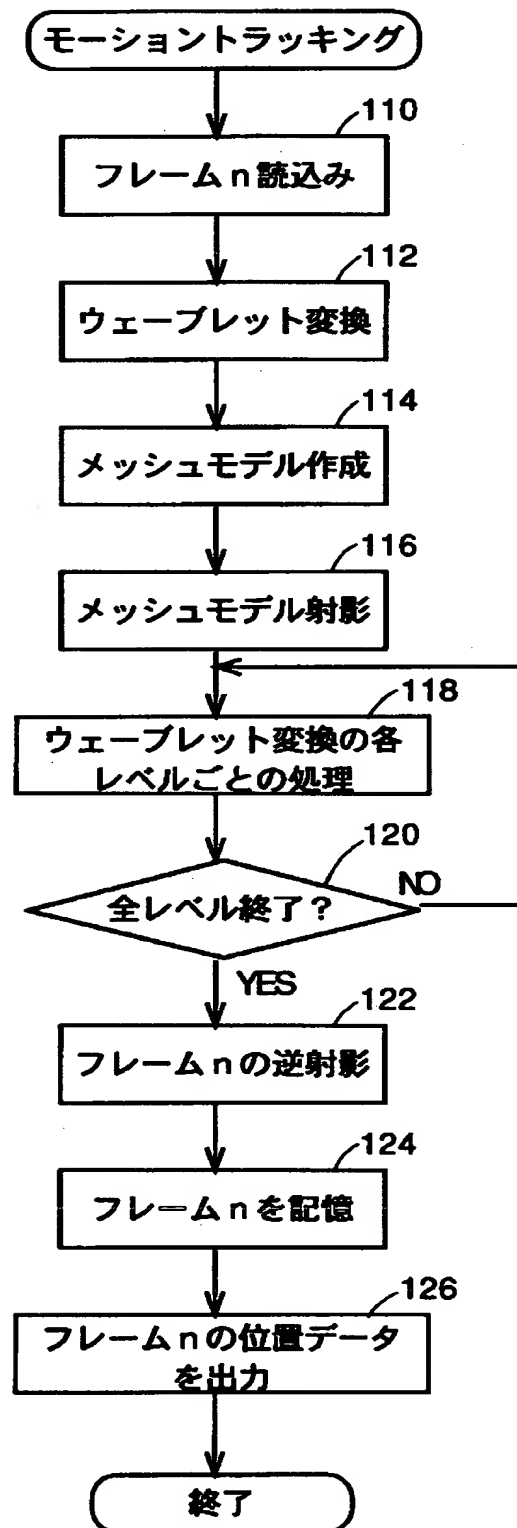


【図 4】



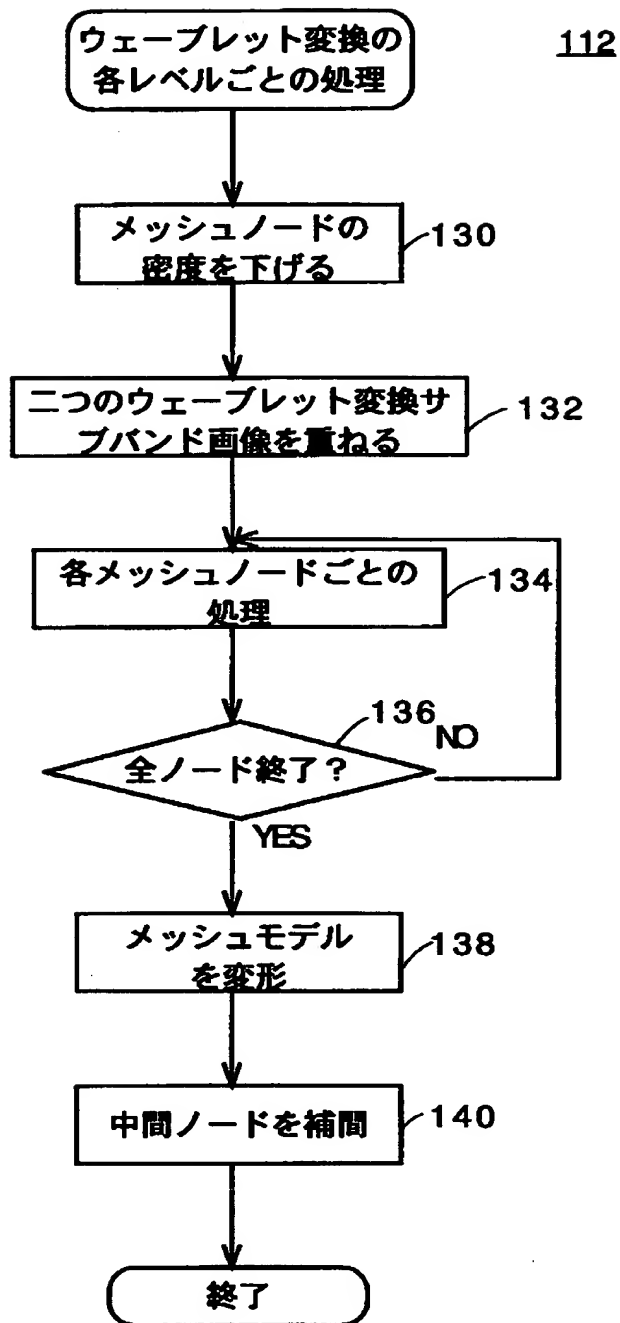
80

【図 5】

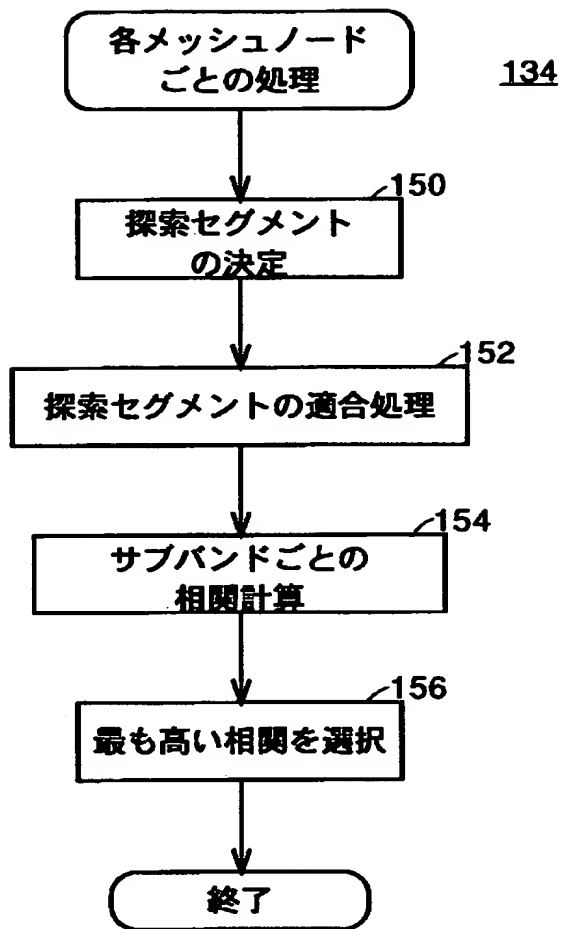


82

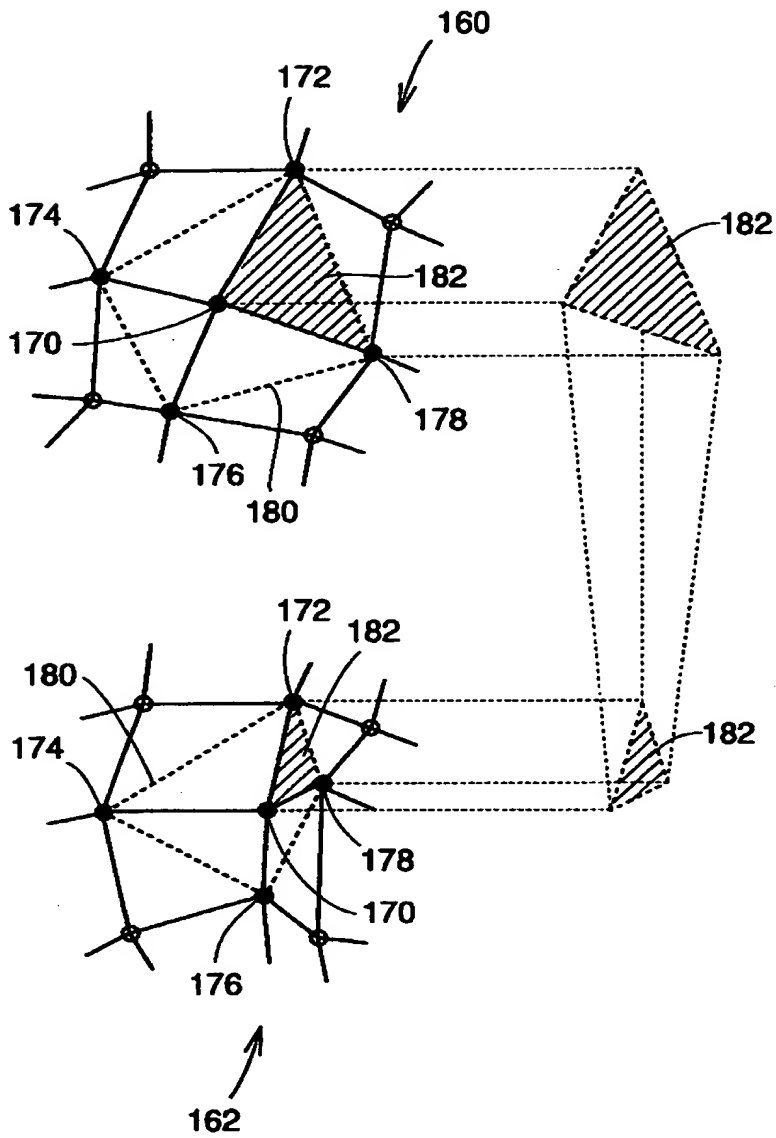
【図 6】



【図 7】



【図 8】



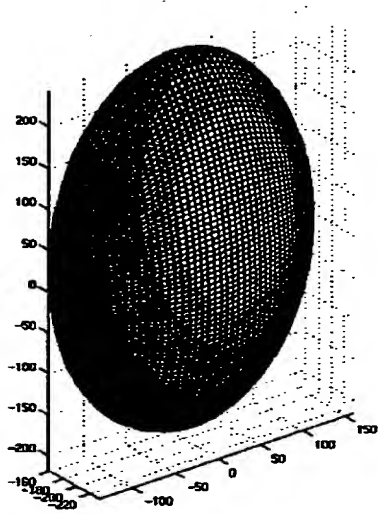
【図 9】



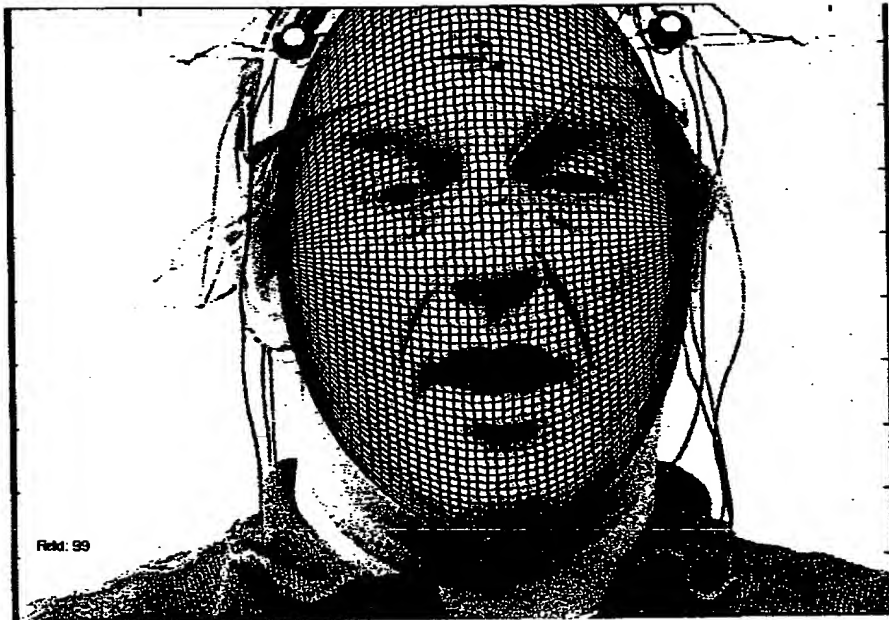
【図 1 0】



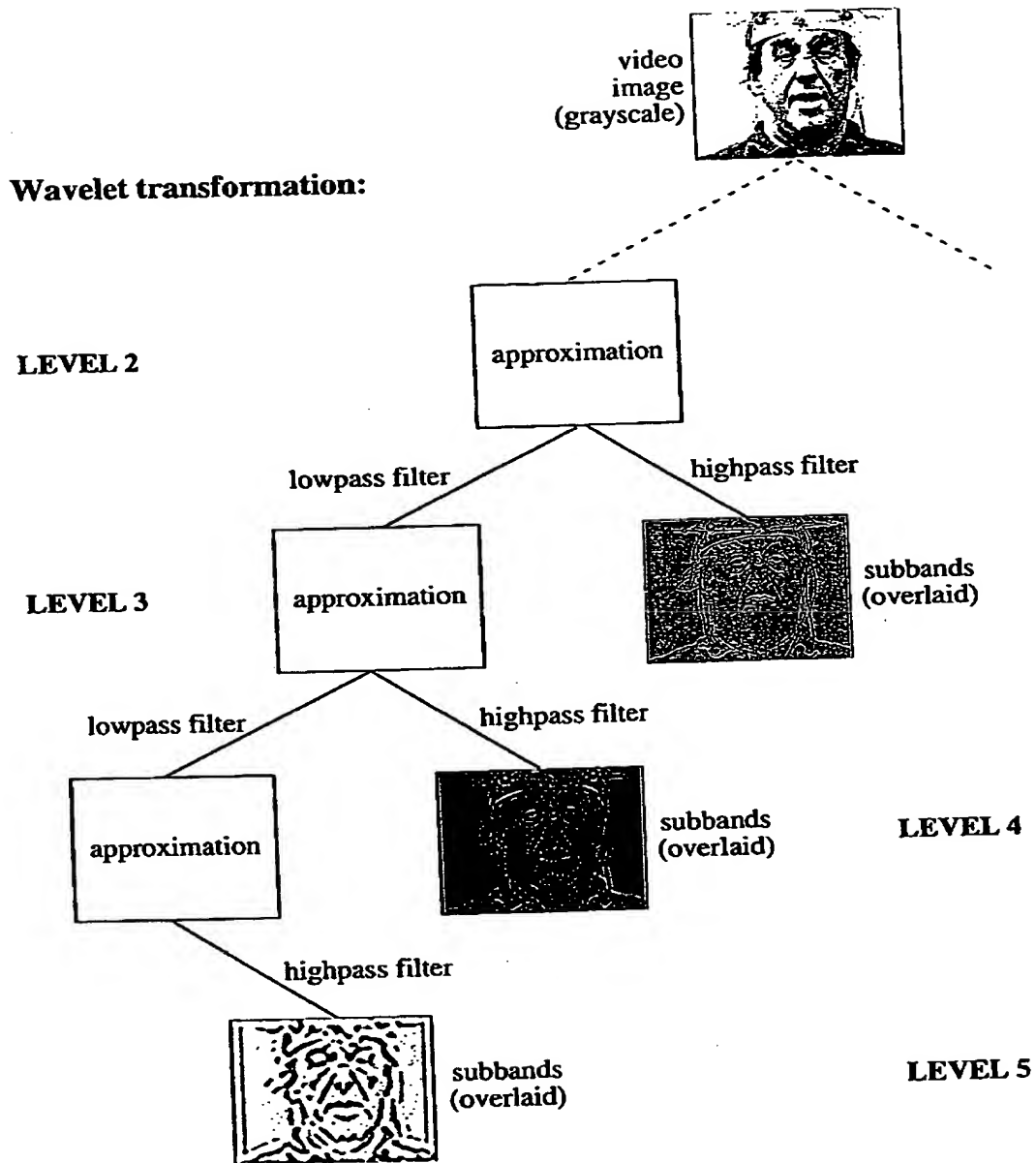
【図 1 1】



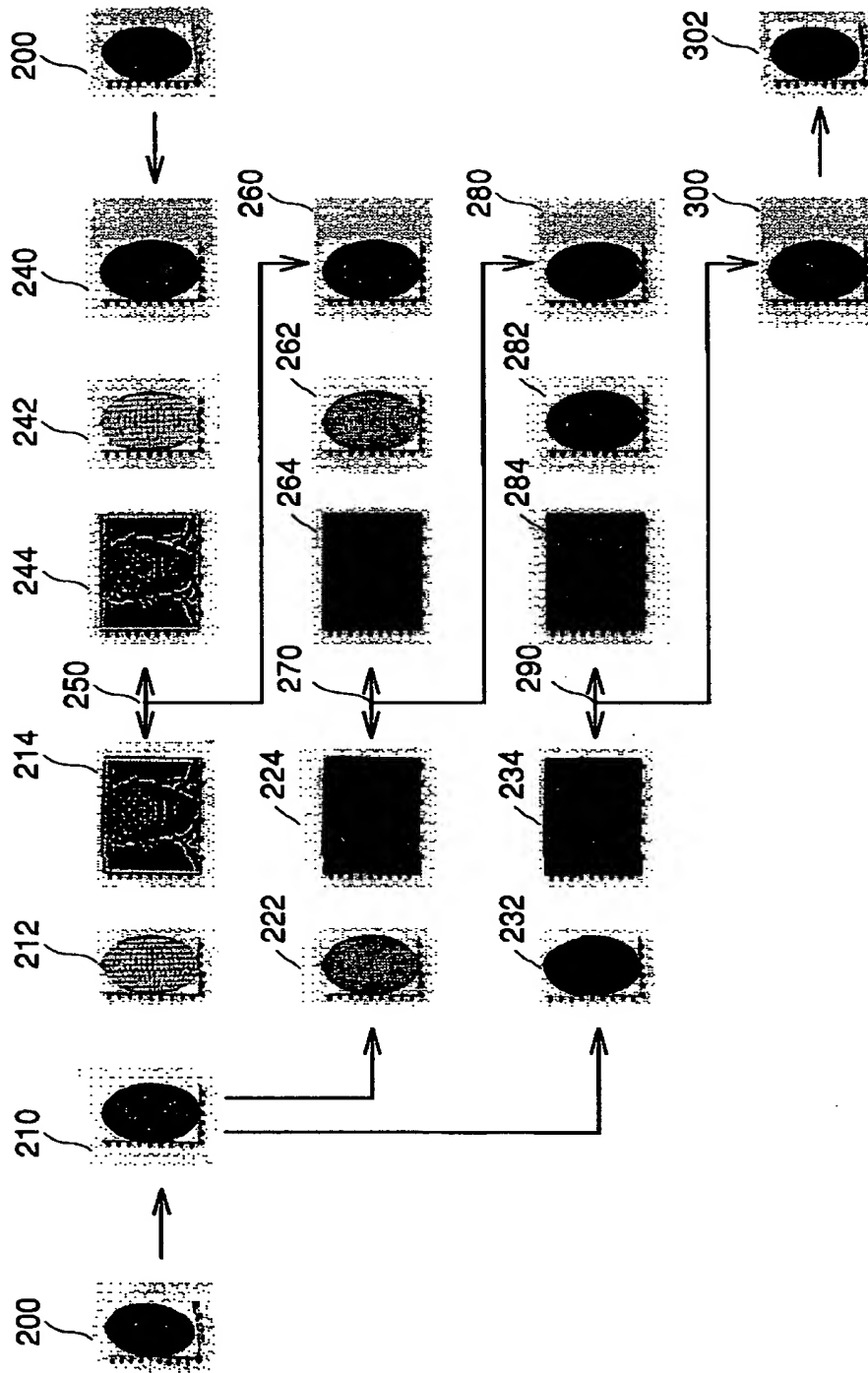
【図 1 2】



【図 1 3】



【図 1 4】



【書類名】 要約書

【要約】

【課題】 ビデオ画像から、信頼性高くかつリアルタイムで人物などの対象物の表面の動きに関する定量的な情報を抽出することができるようにする。

【解決手段】 表面動き測定方法は、メッシュモデルにより第1のフレームの対象物の表面の各部位の位置を表現するステップと、複数通りの解像度で第1および第2のフレームの画像を各々複数のサブバンド画像に変換するステップと、メッシュモデルより低いノード密度の複数の粗メッシュモデルを生成するステップ130と、第1および第2のフレームの各々のサブバンド画像に対して対応する粗メッシュモデルを重ねたオーバーレイ画像を得るステップ132と、第1および第2のフレームの、対応するサブバンド画像からのオーバーレイ画像のテクスチャマップのクロス相関を計算して、各ノードの第1のフレームから第2のフレームの間の相対的位置変化を算出するステップ136とを含む。

【選択図】 図6

出 願 人 履 歴 情 報

識別番号

[592179296]

1. 変更年月日 1992年 8月19日

[変更理由] 新規登録

住 所 京都府相楽郡精華町大字乾谷小字三平谷5番地
氏 名 株式会社エイ・ティ・アール人間情報通信研究所